

# SPEECH COMMUNICATION DEVICE AND ITS COMMUNICATION METHOD

**Patent number:** JP2001184098

**Publication date:** 2001-07-06

**Inventor:** ISHIKAWA TAKAYUKI

**Applicant:** NIPPON-ELECTRIC CO

**Classification:**

- international: **G10L19/00; G10L19/12; H03M7/30; G10L19/00; H03M7/30;**  
(IPC1-7): G10L19/12; H03M7/30

- european:

**Application number:** JP19990365640 19991222

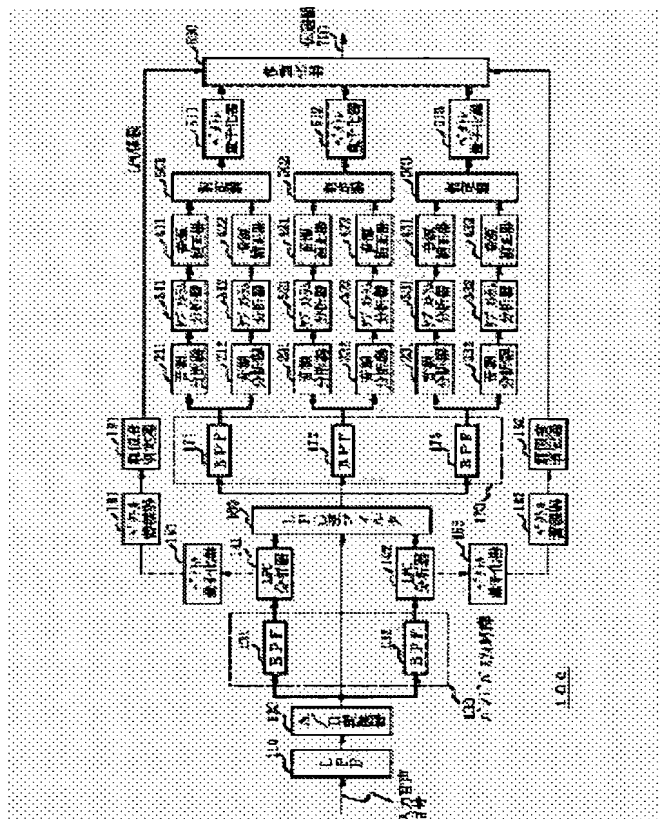
**Priority number(s):** JP19990365640 19991222

Report a data error here

## Abstract of JP2001184098

**PROBLEM TO BE SOLVED:** To permit the more faithful reproduction of a spectrum and narrow-band voice communication with high compressibility of information without increasing an LPC coefficient.

**SOLUTION:** An input voice signal framed by LPC analyzers by two-divided bands is analyzed by linear prediction to find an LPC coefficient and vector quantizers 151 and 152 perform vector quantization. A vector accumulator accumulates the vector-quantized LPC coefficient in large frame units of 10 frames and a similarity decision unit decides mutually similar LPC coefficients among the accumulated, coefficients and further selects only one of them as a representative value. A sound source signal which is processed by an LPC reverse filter, a sound source analyzer, a cepstrum analyzer, a sound source corrector, etc., is quantized by vector quantizers 511 to 513, multiplied together with the mentioned representative vector-quantized LPC coefficient and nonsimilar vector-quantized LPC coefficients, and sent as an encoded voice signal to a transmission line.



Data supplied from the esp@cenet database - Worldwide

# PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2001-184098

(43)Date of publication of application : 06.07.2001

(51)Int.Cl.

G10L 19/12

G10L 19/00

H03M 7/30

(21)Application number : 11-365640

(71)Applicant : NEC CORP

(22)Date of filing : 22.12.1999

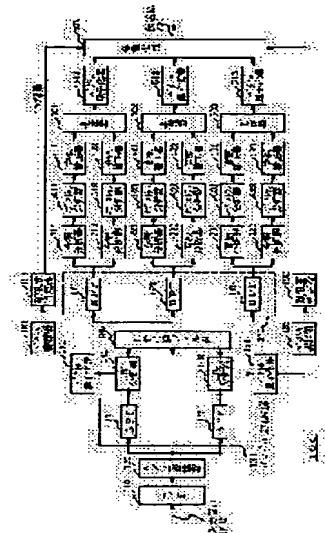
(72)Inventor : ISHIKAWA TAKAYUKI

## (54) SPEECH COMMUNICATION DEVICE AND ITS COMMUNICATION METHOD

### (57)Abstract:

**PROBLEM TO BE SOLVED:** To permit the more faithful reproduction of a spectrum and narrow-band voice communication with high compressibility of information without increasing an LPC coefficient.

**SOLUTION:** An input voice signal framed by LPC analyzers by two-divided bands is analyzed by linear prediction to find an LPC coefficient and vector quantizers 151 and 152 perform vector quantization. A vector accumulator accumulates the vector-quantized LPC coefficient in large frame units of 10 frames and a similarity decision unit decides mutually similar LPC coefficients among the accumulated, coefficients and further selects only one of them as a representative value. A sound source signal which is processed by an LPC reverse filter, a sound source analyzer, a cepstrum analyzer, a sound source corrector, etc., is quantized by vector quantizers 511 to 513, multiplied together with the mentioned representative vector-quantized LPC coefficient and nonsimilar vector-quantized LPC coefficients, and sent as an encoded voice signal to a transmission line.



## LEGAL STATUS

[Date of request for examination]

15.11.2000

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

3468184

[Date of registration]

05.09.2003

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2001-184098

(P2001-184098A)

(43) 公開日 平成13年7月6日 (2001.7.6)

(51) Int.Cl.<sup>7</sup>

識別記号

F I

ターマート\* (参考)

G 1 0 L 19/12

H 0 3 M 7/30

A 5 D 0 4 5

19/00

G 1 0 L 9/14

S 5 J 0 6 4

H 0 3 M 7/30

N

審査請求 有 請求項の数 6 O L (全 10 頁)

(21) 出願番号 特願平11-365640

(22) 出願日 平成11年12月22日 (1999. 12. 22)

(71) 出願人 000004237

日本電気株式会社

東京都港区芝五丁目7番1号

(72) 発明者 石川 孝行

東京都港区芝五丁目7番1号 日本電気株式会社内

(74) 代理人 100082935

弁理士 京本 直樹 (外2名)

Fターム(参考) 5D045 CA01 CB01

5J064 AA02 BA13 BB03 BC06 BC07

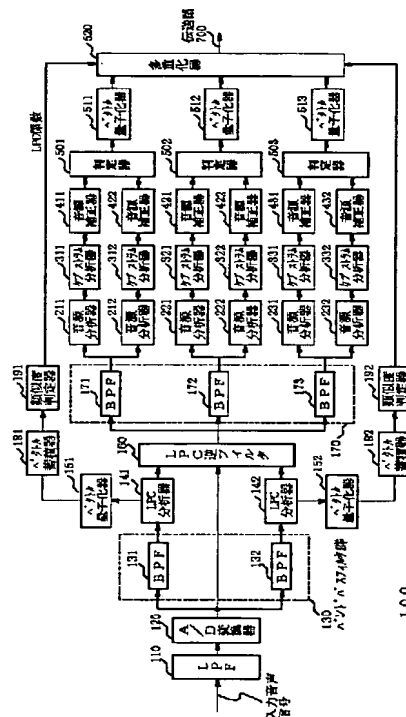
BC11 BC14 BC18 BC25 BD02

(54) 【発明の名称】 音声通信装置及びその通信方法

(57) 【要約】

【課題】 L P C 係数の増加なく、より忠実なスペクトルの再現と、より情報の圧縮度が高い狭帯域音声通信を可能とする。

【解決手段】 2分割帯域ごとの L P C 分析器でフレーム化された入力音声信号を線形予測分析し L P C 係数を求め、ベクトル量子化器 1 5 1, 1 5 2 でベクトル量子化する。ベクトル蓄積器は、10フレーム分集成した大フレーム単位でベクトル量子化 L P C 係数を蓄積し、類似度判定器が、その中から互いに類似性のあるもの判別し、更にその中から一つのみを代表値として選択する。一方、L P C 逆フィルタ、音源分析器、ケプストラム分析器、音源補正器等で処理された音源信号は、ベクトル量子化器 5 1 1 ~ 5 1 3 でベクトル量子化され、多重化器で、前述の代表ベクトル量子化 L P C 係数及び類似性のないベクトル量子化 L P C 係数とともに多重化され、符号化音声信号として伝送路へ送出される。



**【特許請求の範囲】**

**【請求項1】** 予め定めた音声帯域を第1の分割数の帯域に分割し、その分割帯域のそれぞれについて所定周期のフレーム化された入力音声信号を線形予測分析して線形予測係数を出力する線形予測分析手段と、前記線形予測分析手段から出力される分割帯域ごとの前記線形予測係数をベクトル量子化する第1の量子化手段と、同じ分割帯域ごとの前記線形予測係数と前記入力音声信号とを受け、残差信号を抽出する逆フィルタ手段と、前記残差信号を第2の分割数の帯域に分割し、その分割帯域のそれぞれの残差信号に基づいて分割帯域ごとに音源信号を抽出する音源分析手段と、前記分割帯域ごとの音源信号からケプストラム信号を抽出するケプストラム分析手段と、前記ケプストラム信号を基に前記音源信号を補正する音源補正手段と、前記音源補正手段から出力される分割帯域ごとの補正された前記音源信号をベクトル量子化する第2の量子化手段と、前記第1の量子化手段から出力される分割帯域ごとの前記ベクトル量子化済線形予測係数をフレームを所定数分集成した大フレーム単位で蓄積するベクトル蓄積手段と、前記ベクトル蓄積手段に蓄積された複数のベクトル量子化済線形予測係数を互いに類似性のあるものとなないものとに分離し、類似性ありとみなした複数のベクトル量子化済線形予測係数の中から一つのみを代表ベクトル量子化済線形予測係数として選択する類似度判定手段と、前記類似度判定手段から出力される前記代表ベクトル量子化済線形予測係数及び類似性ありとみなされなかった残りのベクトル量子化済線形予測係数と、前記第2の量子化手段から出力される分割帯域ごとの前記ベクトル量子化済音源信号とをそれぞれ多重化し符号化音声信号として出力する多重化手段とを備えることを特徴とする音声通信送信装置。

**【請求項2】** 請求項1記載の音声通信送信装置からの符号化音声信号を受け、その信号から分割帯域ごとの代表ベクトル量子化済線形予測係数及び類似性ありとみなされなかった残りのベクトル量子化済線形予測係数と、当該分割帯域ごとのベクトル量子化済音源信号とをそれぞれ分離し、前記代表ベクトル量子化済線形予測係数を類似性ありとみなしたフレーム数分のベクトル量子化済線形予測係数として複製し、前記類似性ありとみなされなかった残りのベクトル量子化済線形予測係数とともに大フレーム分の値に復元するとともに、各ベクトル値をスカラー値に復元する分離手段と、前記分離手段から出力される分割帯域ごとの前記線形予測係数及び音源信号から全帯域分の合成音声信号を再生する合成手段とを備えることを特徴とする音声通信受信

装置。

**【請求項3】** 音声信号を入力し符号化音声信号として伝送路へ出力する請求項1記載の音声通信送信装置と、前記伝送路を介して前記音声通信送信装置と接続され入力された符号化音声信号から合成音声信号を再生する請求項2記載の音声通信受信装置とから成ることを特徴とする音声通信装置。

**【請求項4】** 予め定めた音声帯域を第1の分割数の帯域に分割し、その分割帯域のそれぞれについて所定周期のフレーム化された入力音声信号を線形予測分析して線形予測係数を求め、前記線形予測分析手段から出力される分割帯域ごとの前記線形予測係数をベクトル量子化し、同じ分割帯域ごとの前記線形予測係数と前記入力音声信号とから残差信号を抽出し、前記残差信号を第2の分割数の帯域に分割し、その分割帯域のそれぞれの残差信号に基づいて分割帯域ごとに音源信号を抽出し、前記分割帯域ごとの音源信号からケプストラム信号を抽出し、前記ケプストラム信号を基に前記音源信号を補正し、前記分割帯域ごとの補正された音源信号をベクトル量子化し、前記分割帯域ごとの前記ベクトル量子化済線形予測係数をフレームを所定数分集成した大フレーム単位で蓄積し、蓄積された複数のベクトル量子化済線形予測係数を互いに類似性のあるものとなないものとに分離し、類似性ありとみなした複数のベクトル量子化済線形予測係数の中から一つのみを代表ベクトル量子化済線形予測係数として選択し、前記代表ベクトル量子化済線形予測係数及び類似性ありとみなされなかった残りのベクトル量子化済線形予測係数と、前記ベクトル量子化済音源信号とをそれぞれ多重化し符号化音声信号として出力することを特徴とする音声通信送信方法。

**【請求項5】** 請求項4記載の音声通信送信方法により処理された符号化音声信号を受信し、その信号から分割帯域ごとの代表ベクトル量子化済線形予測係数及び類似性ありとみなされなかった残りのベクトル量子化済線形予測係数と、当該分割帯域ごとのベクトル量子化済音源信号とをそれぞれ分離し、前記代表ベクトル量子化済線形予測係数を類似性ありとみなしたフレーム数分のベクトル量子化済線形予測係数として複製し、前記類似性ありとみなされなかった残りのベクトル量子化済線形予測係数とともに大フレーム分の値に復元するとともに、各ベクトル値をスカラー値に復元し、前記復元された分割帯域ごとの線形予測係数及び音源信号から全帯域分の合成音声信号を再生することを特徴とする音声通信受信方法。

【請求項6】 送信側にて、入力音声信号を請求項4記載の音声通信送信方法により処理し符号化音声信号として伝送路へ送出し、受信側にて、前記伝送路を介して受信した符号化音声信号を請求項5記載の音声通信受信方法により合成音声信号として再生することを特徴とする音声通信方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は音声通信装置及びその通信方法に関し、特に線形予測符号化(LPC)分析・合成方式を用いた音声通信装置と、その送信装置及び受信装置と、それらの通信方法とに関する。

【0002】

【従来の技術】LPC係数と残差信号を用いたLPC分析・合成方式の音声通信装置として、従来よりピッチ励振型LPCボコーダが知られている。図3はこの従来の音声通信装置(送信装置及び受信装置)としてのピッチ励振型LPCボコーダの一例のブロック図を示す。

【0003】同図において、送信装置側の入力音声信号は音声帯域制限用ローパスフィルタ(LPF)22により、例えば300Hz～3.4kHzの電話音声帯域に制限された後、A/D変換器23に供給されて、所定の標準化周波数で標準化された所定量子化ビット数の音声データに変換される。

【0004】この音声データは、線形予測分析器(LPC分析器)24に供給され、ここで公知の線形予測分析によりkパラメータ、 $\alpha$ パラメータ等の8個～12個程度のLPC係数に変換される。音源分析器25はこのLPC係数を用いて公知の方法で残差信号を抽出し、更にこれを2乗和して音源信号(電力)を算出する。また、上記の音声データは、ピッチ抽出器26に供給されて、音源データのピッチ周波数(声帯振動数)が抽出される。

【0005】以上のLPC分析器24から出力されたLPC係数と、音源分析器25から出力された音源信号(電力)と、ピッチ抽出器26から出力されたピッチ周波数とは、それぞれ多重化器27に供給され、ここで多重化された後、伝送路を介して合成側(受信装置側)へ送出される。

【0006】受信装置側では、分離器28で入力された信号から上記のLPC係数、音源信号(電力)及びピッチ周波数をそれぞれ分離する。パルス列発生器29は分離器28からのピッチ周波数に応じたインパルス列を発生する。有声/無声判定器30は分離器28からのピッチ周波数及び電力により有声音声が無声音声を判定し、その判定結果をスイッチ回路32へスイッチング信号として供給する。

【0007】スイッチ回路32は、有声音声判定時は上記のスイッチング信号に基づき、パルス列発生器29からピッチ周期に対応した固定周期(固定間隔)で取り出

されたインパルス列を選択して音源復調器33へ供給し、無声音声判定時は声帯振動が無くピッチ周波数が存在しないため、前記ピッチ周波数に基づき固定周期的に励振されるインパルス列の代わりに、上記のスイッチング信号に基づき、雑音発生器31から取り出された白色雑音信号に応じたランダムパルス列を選択して音源復調器33へ供給する。

【0008】音源復調器33は、スイッチ回路32からのパルス列と分離器28により分離された電力とに基づいて音源信号を復調する。LPC合成フィルタ34はこの復調音源信号に励振され、分離器28により分離されたLPC係数により係数が制御されてディジタル合成音声信号を出力する。

【0009】このディジタル合成音声信号は、D/A変換器35に供給されてアナログ信号に変換された後、LPF36により不要周波数成分が除去されて電話音声帯域の音声信号とされて出力される。

【0010】このように、従来の音声通信装置では、有声音声の合成に際し、音源情報を電話音声帯域(300Hz～3.4kHz)で一括分析し、その結果得られたピッチ周波数に応じた固定周期でパルス列発生器29からインパルス列を発生している。

【0011】

【発明が解決しようとする課題】しかし、本来の自然音声(有声音声)声帯振動は、一定(固定周期)ではなく、肉声の変化に応じた時間的、周波数的揺らぎ、瞬時変動を伴う準周期的な動作をしている。しかるに、上記の従来の音声通信装置では、音源情報(ピッチ周波数)を分析するにあたって、上述した肉声の周波数的揺らぎ、時間的揺らぎを無視して、電話音声帯域(300Hz～3.4kHz)を定常信号とみなし、全帯域を一括して分析しているため、その肉声の変動に伴う時間的、周波数的な揺らぎを表現し得ず、このため、自然性に乏しい機械的合成音声しか生成できないという問題がある。

【0012】一般に周波数分析の欠点は、その周波数の定倍の信号を検出してしまうことがあることであり、ピッチ周波数分析においても、その倍の周波数(倍ピッチ)や半分の周波数(半ピッチ)を検出して、音声品質の自然性が劣化した合成音声を生成してしまうことがある。

【0013】また、従来、音声信号(残差信号)の帯域内信号を一括して有声・無声と判定しているが、帯域内すべてが同じ状態(有声・無声)でいることは無く、周波数帯域ごとに有声・無声が混在することがある。また、有声音声時においても、各周波数帯域ごとにピッチ周期が異なることがある。

【0014】更に、上記の従来の音声通信装置では、LPC分析は帯域内スペクトルを一括して表現させるため、通常使用する8個～12個程度のLPC係数では、

エネルギーが集中している低域周波数帯に割り当てられ、高域周波数帯の表現精度が不十分になるため、フォルマント帯域幅の過小推定、高次（第3次）フォルマントの近似性に劣り、忠実なスペクトルの再現ができない場合がある。

【0015】更に音源信号は実際には定常的信号ではなく、揺らぎを伴うものであり、周波数帯域ごとにその揺らぎ幅が異なるものである。

【0016】なお、スペクトル精度の向上を図るため、LPC係数を増加させることも考えられるが、LPC係数の増加は、通信情報量の増加にもつながるため、12個を越えるLPC係数の使用は望ましくなく、実際に狭帯域通信を行う音声通信装置においては、12個を越えるLPC係数の使用は現状では困難である。

【0017】本発明は以上の点に鑑みなされたもので、より自然性の高い音声合成を再現するとともに、LPC係数及び音源情報をベクトル量子化により情報量を圧縮し、かつ、相互類似度の高いLPC係数は代表を1つのみ選択し更に情報量を圧縮することにより、LPC係数を増加させることなく（情報量の増加を伴わずに）、より忠実なスペクトルの再現（高品質音声）をすることができる音声通信装置及びその通信方法を提供することを目的とする。

【0018】

【課題を解決するための手段】本発明の音声通信送信装置は；予め定めた音声帯域を第1の分割数の帯域に分割し、その分割帯域のそれぞれについて所定周期のフレーム化された入力音声信号を線形予測分析して線形予測係数を出力する線形予測分析手段と；前記線形予測分析手段から出力される分割帯域ごとの前記線形予測係数をベクトル量子化する第1の量子化手段と；同じ分割帯域ごとの前記線形予測係数と前記入力音声信号とを受け、残差信号を抽出する逆フィルタ手段と；前記残差信号を第2の分割数の帯域に分割し、その分割帯域のそれぞれの残差信号に基づいて分割帯域ごとに音源信号を抽出する音源分析手段と；前記分割帯域ごとの音源信号からケプストラム信号を抽出するケプストラム分析手段と；前記ケプストラム信号を基に前記音源信号を補正する音源補正手段と；前記音源補正手段から出力される分割帯域ごとの補正された前記音源信号をベクトル量子化する第2の量子化手段と；前記第1の量子化手段から出力される分割帯域ごとの前記ベクトル量子化済線形予測係数をフレームを所定数分集成した大フレーム単位で蓄積するベクトル蓄積手段と；前記ベクトル蓄積手段に蓄積された複数のベクトル量子化済線形予測係数を互いに類似性のあるものとなし分離し、類似性ありとみなした複数のベクトル量子化済線形予測係数の中から一つのみを代表ベクトル量子化済線形予測係数として選択する類似度判定手段と；前記類似度判定手段から出力される前記代表ベクトル量子化済線形予測係数及び類似性ありと

みなされなかった残りのベクトル量子化済線形予測係数と、前記第2の量子化手段から出力される分割帯域ごとの前記ベクトル量子化済音源信号とをそれぞれ多重化し符号化音声信号として出力する多重化手段とを備える。

【0019】本発明の音声通信受信装置は；上記音声通信送信装置からの符号化音声信号を受け、その信号から分割帯域ごとの代表ベクトル量子化済線形予測係数及び類似性ありとみなされなかった残りのベクトル量子化済線形予測係数と、当該分割帯域ごとのベクトル量子化済音源信号とをそれぞれ分離し、前記代表ベクトル量子化済線形予測係数を類似性ありとみなしたフレーム数分のベクトル量子化済線形予測係数として複製し、前記類似性ありとみなされなかった残りのベクトル量子化済線形予測係数とともに大フレーム分の値に復元するとともに、各ベクトル値をスカラー値に復元する分離手段と；前記分離手段から出力される分割帯域ごとの前記線形予測係数及び音源信号から全帯域分の合成音声信号を再生する合成手段とを備える。

【0020】本発明の音声通信装置は、音声信号を入力し符号化音声信号として伝送路へ出力する上記構成の音声通信送信装置と、前記伝送路を介して前記音声通信送信装置と接続され入力された符号化音声信号から合成音声信号を再生する上記構成の音声通信受信装置とから成る。

【0021】本発明の音声通信送信方法は；予め定めた音声帯域を第1の分割数の帯域に分割し、その分割帯域のそれぞれについて所定周期のフレーム化された入力音声信号を線形予測分析して線形予測係数を求め；前記線形予測分析手段から出力される分割帯域ごとの前記線形予測係数をベクトル量子化し；同じ分割帯域ごとの前記線形予測係数と前記入力音声信号とから残差信号を抽出し；前記残差信号を第2の分割数の帯域に分割し、その分割帯域のそれぞれの残差信号に基づいて分割帯域ごとに音源信号を抽出し；前記分割帯域ごとの音源信号からケプストラム信号を抽出し；前記ケプストラム信号を基に前記音源信号を補正し；前記分割帯域ごとの補正された音源信号をベクトル量子化し；前記分割帯域ごとの前記ベクトル量子化済線形予測係数をフレームを所定数分集成した大フレーム単位で蓄積し；蓄積された複数のベクトル量子化済線形予測係数を互いに類似性のあるものとなし分離し、類似性ありとみなした複数のベクトル量子化済線形予測係数の中から一つのみを代表ベクトル量子化済線形予測係数として選択し；前記代表ベクトル量子化済線形予測係数及び類似性ありとみなされなかった残りのベクトル量子化済線形予測係数と、前記ベクトル量子化済音源信号とをそれぞれ多重化し符号化音声信号として出力する工程を有する。

【0022】本発明の音声通信受信方法は、上記音声通信送信方法により処理された符号化音声信号を受信し、その信号から分割帯域ごとの代表ベクトル量子化済線形

予測係数及び類似性ありとみなされなかった残りのベクトル量子化済線形予測係数と、当該分割帯域ごとのベクトル量子化済音源信号とをそれぞれ分離し、前記代表ベクトル量子化済線形予測係数を類似性ありとみなしたフレーム数分のベクトル量子化済線形予測係数として複製し、前記類似性ありとみなされなかった残りのベクトル量子化済線形予測係数とともに大フレーム分の値に復元するとともに、各ベクトル値をスカラー値に復元し、前記復元された分割帯域ごとの線形予測係数及び音源信号から全帯域分の合成音声信号を再生する工程を有する。

【0023】本発明の音声通信方法は、送信側にて、入力音声信号を上記音声通信送信方法により処理し符号化音声信号として伝送路へ送出し、受信側にて、前記伝送路を介して受信した符号化音声信号を上記音声通信受信方法により合成音声信号として再生する工程を有する。

【0024】

【発明の実施の形態】まず、本発明の概要を説明する。本発明の音声通信装置は、音声信号をスペクトラル包絡情報及び音源信号を生成し符号化音声信号として伝送する音声通信装置に関し、複数種類の周波数分析器、逆フィルタ、音源補正器、及びベクトル量子化器を備えた音声通信送信装置と、この音声通信送信装置と伝送路を介して接続される音源信号復調器を備えた音声通信受信装置とから成る。

【0025】音声通信送信装置は、音声帯域を第1の分割帯域それぞれについて所定周期のフレーム化された入力音声信号を線形予測分析(LPC分析)して線形予測係数(LPC係数)を出力する線形予測分析手段と、線形予測係数をベクトル量子化する第1の量子化手段と、線形予測係数と入力音声信号との残差信号を抽出する逆フィルタ手段と、残差信号から第2の分割帯域ごとに音源信号を抽出する音源分析手段と、音源信号からケプストラム信号を抽出するケプストラム分析手段と、ケプストラム信号を基に音源信号を補正する音源補正手段と、補正された音源信号をベクトル量子化する第2の量子化手段と、ベクトル量子化済線形予測係数をフレームを所定数分集成した大フレーム単位で蓄積するベクトル蓄積手段と、蓄積された複数のベクトル量子化済線形予測係数を互いに類似性のあるものとなないものとに分離し、類似性ありとみなした複数のベクトル量子化済線形予測係数の中から一つのみを代表ベクトル量子化済線形予測係数として選択する類似度判定手段と、ベクトル量子化済線形予測係数及び類似性ありとみなされなかった残りのベクトル量子化済線形予測係数と、ベクトル量子化済音源信号とをそれぞれ多重化し符号化音声信号として出力する多重化手段とを備える。

【0026】音声通信受信装置は、符号化音声信号を受け、その信号から分割帯域ごとの代表ベクトル量子化済線形予測係数及び類似性ありとみなされなかった残りのベクトル量子化済線形予測係数と、当該分割帯域ごとの

ベクトル量子化済音源信号とをそれぞれ分離し、代表ベクトル量子化済線形予測係数を類似性ありとみなしたフレーム数分のベクトル量子化済線形予測係数として複製し、類似性ありとみなされなかった残りのベクトル量子化済線形予測係数とともに大フレーム分の値に復元するとともに、各ベクトル値をスカラー値に復元する分離手段と、線形予測係数及び音源信号から合成音声信号を再生する合成手段とを備える。

【0027】本発明では、分割帯域のそれぞれについて入力音声信号を線形予測分析して線形予測係数(LPC係数)を得るようにしたため、従来のLPC分析の欠点といわれていた(1)フォルマント帯域幅の過小推定、(2)第3フォルマントの近似性の悪さをそれぞれ改善でき、このLPC係数を用いて逆フィルタにより高精度の残差信号を抽出することができる。

【0028】また、この高精度の残差信号を更に帯域分割し、それぞれについて音源分析手段にて音源信号を得るようにしたため、各帯域ごとに混在している有声音声/無声音声を分離し、有声音声においては各帯域ごとに最適な音源情報の抽出ができる。

【0029】更に、各帯域ごとに音源信号の揺らぎを測定し、かつ補正するため、自然音声に近い揺らぎを持つ音源信号の抽出ができる。従って、個々の特性に合わせた精度の良い音源情報を抽出することができるため、本来の人間の発声機構に忠実な音声分析・合成ができる。

【0030】加えて、ベクトル量子化済線形予測係数を全て伝送するのではなく、類似性のあるベクトル値は一つの代表ベクトル値のみ伝送する代表ベクトル送出方式であるため、音質を保ちながら、通信情報量の削減をすることができる。

【0031】次に、本発明の実施の形態について図面を参照して詳細に説明する。

【0032】図1は本発明の音声通信送信装置100の一実施の形態を示すブロック図である。同図において、入力音声信号は音声帯域制限用ローパスフィルタ(LPF)110により、例えば300Hz～3.4kHzの電話音声帯域に制限された後、アナログ・デジタル変換器(A/D変換器)120に供給されて、所定の標準化周波数(例えば8kHz)で標準化された所定量子化ビット数(例えば16ビット)のデジタル音声データに変換される。この音声データは所定周期(本例では2.5msec)で連続するフレームで構成されており、以下の処理においては、このフレーム単位で処理される。

【0033】この音声データは、バンドパスフィルタ群130に供給されて、所定の周波数帯域ごとの複数の帯域に分割される。本例では、1.2kHzを中心に2分割される。すなわち、音声データは、バンドパスフィルタ群130を構成するバンドパスフィルタ(BPF)131及び132にそれぞれ供給され、BPF131から

帯域が300Hz～1.2kHz、BPF132から帯域が1.2kHz～3.4kHzとされて取り出される。

【0034】BPF131及び132から出力された各帯域分割音声データは、分割帯域ごとに対応して設けられた線形予測分析器(LPC分析器)141及び142にそれぞれ供給され、公知の線形予測分析によりLSPパラメータ、 $\alpha$ パラメータ等のLPC係数(線形予測係数)に変換された後、LPC逆フィルタ160にA/D変換器120の出力音声データと共に入力される。

【0035】各LPC分析器141及び142からのLPC係数(本例では $\alpha$ パラメータを使用)を圧縮・量子化するためのベクトル量子化器151及び152にそれぞれ入力し、各々のLPC係数のベクトル量子化値(ベクトル量子化済LPC係数)を算出しベクトル蓄積器181及び182にそれぞれ送出する。

【0036】各ベクトル蓄積器181及び182は、音声の声道周期として違和感のない100msec～220msec程度の時間長に相当するように音声データのフレームを複数集成して大フレームとし(本例は1フレーム22.5msecとして、10フレームを1大フレームとしている)、この大フレーム単位にそれを構成する各フレームに対応するベクトル値(ベクトル量子化済LPC係数)を蓄積するとともに、類似度判定器191及び192にそれぞれ送出する。

【0037】各類似度判定器191及び192は、前述の大フレームを構成する10フレーム分のベクトル信号(ベクトル量子化済LPC係数)の中から互いに類似性のある信号を抽出し、その中の一つのみを代表ベクトルとして選択する。実験的には、有声音声の場合は10フレーム中2～5フレームが類似性があると算出される場合が多い。よって線形予測ベクトル量子化情報量の2～5フレーム分のデータを1フレーム分のデータとして選択された代表ベクトル値(代表ベクトル量子化済LPC係数)と、類似性なしと判定されたベクトル値(ベクトル量子化済LPC係数)とを多重化器10へ送る。このように、互いに類似性のある2～5フレーム分のデータを1フレーム分のデータとするので1/2～1/5に情報圧縮が可能である。

【0038】一方、LPC逆フィルタ160は、A/D変換器120の出力音声データに対し、LPC分析器141及び142からの2つの分割帯域のLPC係数を用いて線形予測分析のスペクトル包絡特性とは逆のフィルタ特性を付与し、残差信号を出力する。

【0039】この残差信号は、バンドパスフィルタ群170に供給されて複数の帯域に分割される。本例では、1.5kHzと2.5kHzを境にして帯域が3分割される。すなわち、残差信号は、バンドパスフィルタ群170を構成するBPF171、172、及び173にそれぞれ供給され、BPF171から300Hz～1.5

kHzの周波数成分が取り出され、BPF172から1.5kHz～2.5kHzの周波数成分が取り出され、BPF173から2.5kHz～3.4kHzの周波数成分が取り出される。

【0040】各BPF171、172、及び173からそれぞれ取り出された帯域分割残差信号は、分帯域ごとに対応してそれぞれ複数個(本例は2個)ずつ設けられた音源分析器211～212、221～222、及び231～232にそれぞれ供給され、ここで2乗和されて分割帯域ごとの音源信号(電力)として取り出される。

【0041】各音源分析器211～212、221～222、及び231～232からの各帯域ごとの音源信号は、対応するケプストラム分析器311～312、321～322、及び331～332にそれぞれ供給される。

【0042】各ケプストラム分析器311～312、321～322、及び331～332は、各帯域ごとの音源信号の周波数軸上のケプストラム値(ケフレンシー)を算出し、対応する音源補正器411～412、421～422、及び431～432にそれぞれ出力する。

【0043】各音源補正器411～412、421～422、及び431～432は、ケプストラム値のケフレンシー上の分散値を測定し、分散が小さいときには、定常信号と判断、分散が大きいたまきは揺らぎがある信号と判断する。揺らぎがあると判断された音源信号はピッチ周期(声帯信号)に同期した変動(揺らぎ)を与え補正する。更に、各音源補正器411～412、421～422、及び431～432の出力は、次に説明する各帯域対応の判定器501、502、及び503を通して、第2のベクトル量子化器である、ベクトル量子化器511、512、及び513にそれぞれ出力する。

【0044】なお、音源分析器、ケプストラム分析器、及び音源補正器を各帯域ごとに2つつ設けるのは次の理由による。音源分析にあたり、BPF群170で各帯域に分割された出力信号を、それぞれの帯域ごとの倍ピッチ又は半ピッチを含めて算出するように更に2分割し、それぞれの計算結果からどちらかの分割側(2個のうち1個)がより正しいかを各帯域ごとに判定器501、502、及び503で判定して、入力信号により近い値を示す分析ルート正しい値とし、そのルートから得られた値を正しい各帯域ごとの分析情報とするためである。

【0045】各ベクトル量子化器511、512、及び513は、判定器501、502、及び503でそれぞれ選択された側の音源補正器411～412、421～422、及び431～432で補正された音源信号のベクトル量子化値(ベクトル量子化済音源信号)をそれぞれ算出する。

【0046】多重化器520は、ベクトル量子化器511、512、及び513からそれぞれ取り出された分割



帯域ごとのベクトル量子化された音源信号（電力）と、もう一つのベクトル量子化器151及び152（ベクトル蓄積器、類似度判定器経由）からそれぞれ取り出された2つの分割帯域のベクトル量子化されたLPC係数とをそれぞれ多重化し、符号化音声信号として伝送路700へ出力する。

【0047】図2は本発明の音声通信受信装置800の一実施の形態を示すブロック図である。同図において、分離器810は、伝送路700を介して音声通信送信装置100から入力された符号化音声信号より、送信側と同様の3つの分割帯域のベクトル量子化された音源信号（電力）と2つの分割帯域のベクトル量子化されたLPC係数とをそれぞれ分離するとともに、選択された代表ベクトル値（ベクトル量子化済LPC係数）を類似性ありとみなしたフレーム数分のベクトル値として複製し、類似性なしと判定されたベクトル値とともに大フレーム（10フレーム）分の値に復元した後、通常の線形計算に扱えるスカラー値に復元する。

【0048】このうちLPC係数はLPC補間器820に供給され、音声帯域の上位側と下位側をそれぞれ示しているLPC値（ベクトル値）を、線形演算で扱えるスカラー値に変更した後、両LPC係数（スカラー値）を単純に重ね合わせることで、音声全帯域を表現しているLPC係数に再生し、ここで一定周期（例えば22.5 msec）ごとに入力されるLPC係数の前回の入力値と今回の入力値とを利用した線形補間値から、例えば5.625 msec単位のLPC係数に修正される（換言すると、22.5 msec単位で変化するLPC係数が、5.625 msec単位で変化するLPC係数に変換される）。

【0049】また、分離器810により分離された分割帯域ごとの音源信号（電力）は、音源復調器830に供給され、ここで帯域ごとの音源信号（電力）を補間処理して全帯域（300 Hz～3.4 kHz）のピッチ情報に復元される。ここで音源復調器830は、3つのバンドパスフィルタ群を持ち、それぞれ、低域、中域、高域の音源再生用として、各帯域ごとの音源情報（スカラー値）をフィルタ係数とし、その帯域に対応したピッチ情報をエネルギーとして駆動する。そしてこの3つのフィルタの出力の線形和が、音声全帯域を表す音源情報となる。

【0050】LPC合成フィルタ840は、LPC補間器820から出力された修正後のLPC係数をフィルタ係数とし、音源復調器830から出力された復調後のピッチ情報を入力エネルギーとしたデジタル合成音声データを再生する。

【0051】このデジタル合成音声データは、デジタル・アナログ変換器（D/A変換器）850に供給されてデジタル・アナログ変換されてアナログ信号の合成音声信号として取り出され、次段のLPF860によ

り不要周波数成分を除去された後、再合成音声信号として出力される。

【0052】このように、この実施の形態では、音声信号帯域を2分割して得られたLPC係数を用いてLPC逆フィルタ160により抽出した高精度の残差信号を更に3つの帯域に分割し、それぞれについて音源分析器211～212、221～222、及び231～232にて音源信号を得ることにより、各帯域ごとに混在している有声音声/無声音声を分離し、個々の特性に合わせた精度の良い音源情報を抽出することができ、かつ各帯域ごとに音源信号の揺らぎを測定し補正するため、自然音声に近い揺らぎを持つ音源信号の抽出ができる。従って、本来の人間の発声機構に忠実な音声分析・合成ができる。

【0053】更に周波数分析上の欠点である倍周期検出を補正するために、音源分析器を2分割し、BPF群170で分割されたそれぞれの帯域における倍ピッチ又は、半ピッチを含めて算出しているため、分析エラーを極限できるという特徴を持つ。

【0054】また、声道情報であるベクトル量子化済線形予測（LPC）係数を全て送るのではなく、類似性のあるものは一つの代表値のみを送るため、音質の低下なしで、通信情報量の削減ができる。

【0055】なお、本発明は上記の実施の形態に限定されるものではなく、例えばバンドパスフィルタ群130及び170の各BPFの特性は固定として説明したが、入力情報を基に中心周波数を可変することもでき、また、分割数も2分割から4分割程度まで可能であること、ケプストラムの分散値の判定は本実施例では固定として説明したが、ピッチ周波数に応じて可変してもよいことが実験的に確認されている。従来のLPCボコーダはLPC分析に40ビット、ピッチに7ビット、音源に6ビット、同期に1ビットの計54ビットを1フレーム22.5 msecとして伝送（54/22.5 ms=2400 bps）としている。

【0056】今回の方式はLPC分析におけるベクトル量子化器（151、152）のそれぞれのビット割当を10ビットとし、計20ビットをLPC係数に割り当てる。また、代表値選択情報用としての類似度判定情報を3ビットとし、計6ビットを割り当てる。更に音源情報にしてもベクトル量子化器（511、512、513）のそれぞれのビットを4ビットとし、計12ビットを音源情報に割り当てる。また、ピッチ情報として各帯域の音源ごとに5ビットの計15ビットを割り当てる。（LPC係数20ビット、類似度判定情報6ビット、音源情報12ビット、ピッチ情報15ビット、同期1ビットの計54ビットで実現できる）よって、本発明の実施例において、更なる音質向上をビット情報の増加無く実現できることがわかる。

【0057】

【発明の効果】以上説明したように、本発明によれば、分割帯域のそれぞれについて入力音声信号を線形予測分析して線形予測係数（LPC係数）を用いて逆フィルタにより高精度の残差信号を抽出した後、この高精度の残差信号を更に帯域分割し、それぞれについて音源分析手段にて音源信号を得ることにより、各帯域ごとに混存している有声音声／無声音声を分離するようにしたため、個々の特性に合わせた（各周波数帯域ごとに最適な）精度の良い音源情報を抽出することができるとともに、周波数分析上の欠点といえる、倍ピッチ、半ピッチによる分析エラーを防ぐことができ、各帯域ごとに音源信号の揺らぎを測定しかつ補正するため自然音声に近い揺らぎを持つ音源信号の抽出をし、本来の人間の発声機構に忠実な音声分析・合成ができ、より自然性の高い合成音声を得ることができる。また、ベクトル量子化済線形予測（LPC）係数を全て送るのではなく、類似性のあるものは一つの代表値のみを送るため、LPC係数の増加なく（通信情報量の削減）、より忠実なスペクトルの再現（音質の低下なし）ができるとともに、より情報の圧縮度が高い狭帯域音声通信を行うことができる。

【図面の簡単な説明】

【図1】本発明の音声通信送信装置の一実施の形態を示すブロック図である。

【図2】本発明の音声通信受信装置の一実施の形態を示すブロック図である。

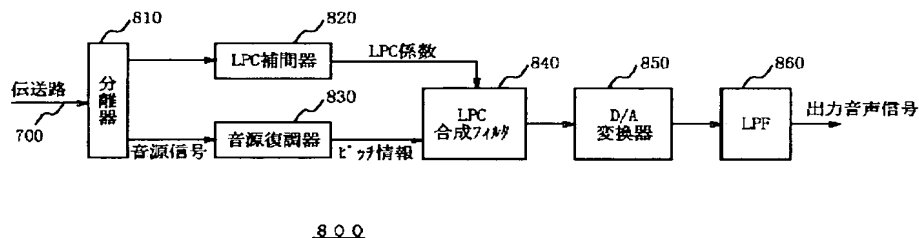
【図3】従来の音声通信送信装置及び音声通信受信装置

を示すブロック図である。

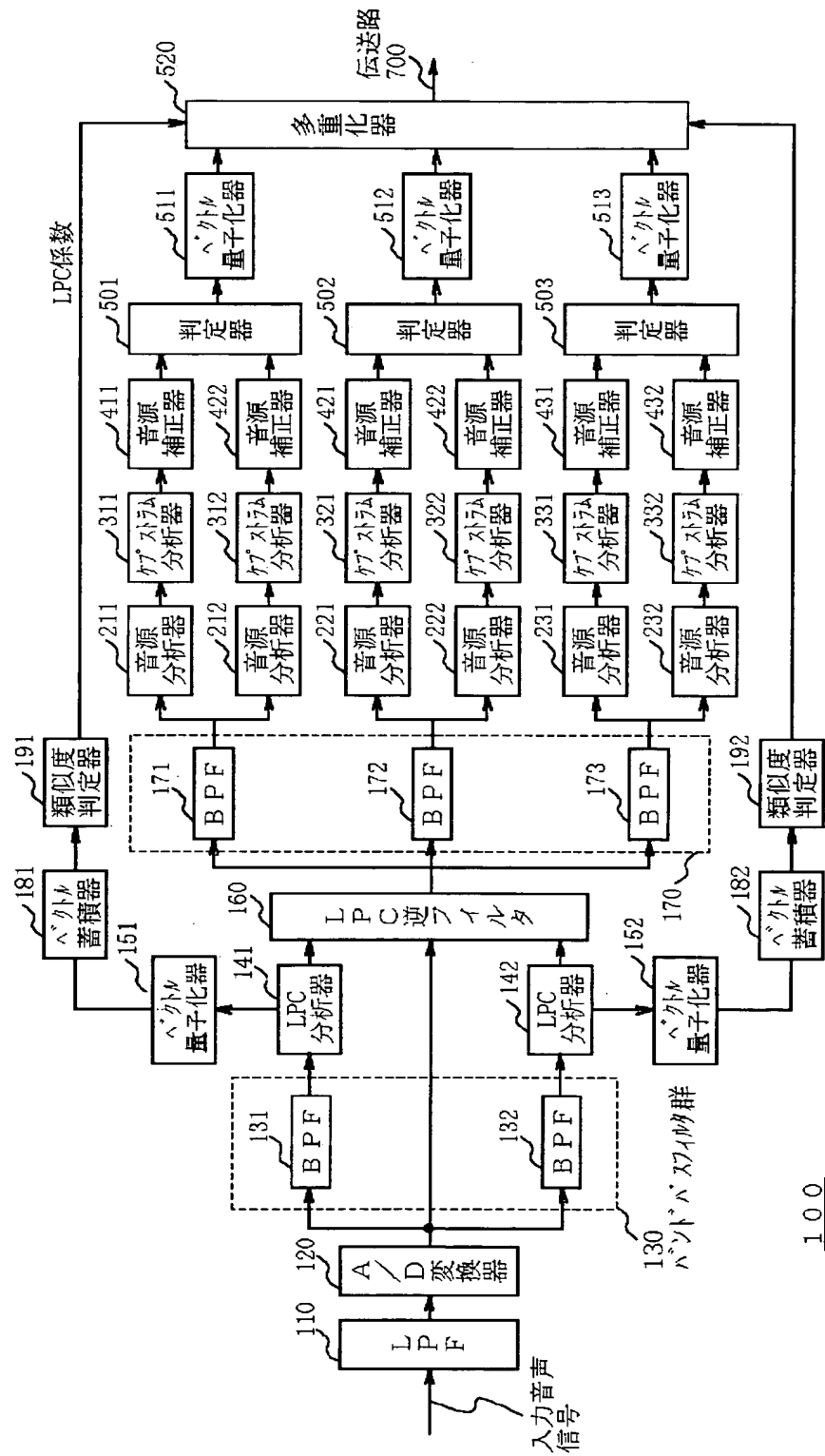
【符号の説明】

100 音声通信送信装置  
 110, 860 ローパスフィルタ（LPF）  
 120 A/D変換器  
 130, 170 バンドパスフィルタ群  
 131, 132, 171～173 バンドパスフィルタ（BPF）  
 141, 142 LPC分析器  
 151, 152, 511～513 ベクトル量子化器  
 160 LPC逆フィルタ  
 211, 212, 221, 222, 231, 232 音源分析器  
 311, 312, 321, 322, 331, 332 ケプストラム分析器  
 411, 412, 421, 422, 431, 432 音源補正器  
 501～503 判定器  
 520 多重化器  
 800 音声通信受信装置  
 810 分離器  
 820 LPC補間器  
 830 音源復調器  
 840 LPC合成フィルタ  
 850 D/A変換器  
 860

【図2】



【図1】



【図3】

